

引用格式:王培晓,张恒才,张彤,等.基于注意力常微分方程的PM<sub>2.5</sub>浓度预测及其可解释性分析[J].地球信息科学学报,2024,26(6):1363-1373. [ Wang P X, Zhang H C, Zhang T, et al. An attentional ordinary differential equation for predicting PM<sub>2.5</sub> concentration and its interpretability analysis[J]. Journal of Geo-information Science, 2024,26(6):1363-1373. ] DOI:10.12082/dqxxkx.2024.230678

# 基于注意力常微分方程的PM<sub>2.5</sub>浓度预测及其可解释性分析

王培晓<sup>1</sup>,张恒才<sup>1,2\*</sup>,张彤<sup>3</sup>,陆锋<sup>1,2,4</sup>

1. 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室,北京 100101; 2. 中国科学院大学资源与环境学院,北京 100049; 3. 武汉大学 测绘遥感信息工程国家重点实验室,武汉 430079; 4. 政务大数据应用省部共建协同创新中心,福州 350003

## An Attentional Ordinary Differential Equation for Predicting PM<sub>2.5</sub> Concentration and Its Interpretability Analysis

WANG Peixiao<sup>1</sup>, ZHANG Hengcai<sup>1,2\*</sup>, ZHANG Tong<sup>3</sup>, LU Feng<sup>1,2,4</sup>

1. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; 2. College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China; 3. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing Science, Wuhan University, Wuhan 430079, China; 4. Fujian Collaborative Innovation Center for Big Data Applications in Governments, Fuzhou 350003, China

**Abstract:** Accurate and explainable prediction of PM<sub>2.5</sub> concentration can help humans avoid exposure risks to air pollution, which is of great significance for human health risk assessment and policy implementation. Currently, the existing PM<sub>2.5</sub> concentration prediction models focus on improving the model prediction accuracy without considering model interpretability, resulting in poor model reusability and trustworthiness. Therefore, this paper proposes an Attentional Spatiotemporal Ordinary Differential Equation (ASTODE) model for PM<sub>2.5</sub> concentration prediction tasks considering both prediction accuracy and model interpretability. Specifically, this

收稿日期 2023-11-15;修回日期 2024-01-23.

基金项目:国家重点研发计划项目(2022YFB3904102);国家博士后创新人才支持计划项目(BX20230360);中国博士后面上资助项目(2023M743454);中国科学院特别研究助理项目;国家自然科学基金项目(42371470);武汉大学测绘遥感信息工程国家重点实验室开放基金项目(23I03);资源与环境信息系统国家重点实验室创新项目(08R8A092YA)。[ **Foundation items:** National Key Research and Development Program of China, No.2022YFB3904102; China National Postdoctoral Program for Innovative Talents, No.BX20230360; China Postdoctoral Science Foundation, No.2023M743454; Special Research Assistant Program of Chinese Academy of Sciences; National Natural Science Foundation of China, No.42371470; Open Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, No.23I03; Innovation Project of LREIS, No.08R8A092YA.]

作者简介:王培晓(1994—),男,山东济南人,博士,研究方向为时空数据分析、时空预测、轨迹分析等。

E-mail: wpx@lreis.ac.cn

\*通讯作者 张恒才(1985—),男,山东济南人,博士,副研究员,研究方向为三维时空智能、城市计算、人类活动大数据分析等。

E-mail: zhanghc@lreis.ac.cn

paper integrates the Neural Ordinary Differential Equation (NODE) into the PM<sub>2.5</sub> concentration prediction task to improve the interpretability of the prediction model. In addition, to address the challenge of traditional NODE in mining spatial dependencies in PM<sub>2.5</sub> concentration data, this paper proposes a novel spatiotemporal derivative network that extends the traditional NODE to spatiotemporal ordinary differential equations. To address the challenges of traditional NODE in mining long-term dependencies in PM<sub>2.5</sub> concentration data, this paper proposes a spatiotemporal attention mechanism to fuse hidden states of multiple time nodes. In the experimental section, the proposed ASTODE model is validated using a real PM<sub>2.5</sub> concentration dataset. This paper quantifies the prediction errors of the ASTODE model in both temporal and spatial dimensions. By comparing with six existing baseline methods, our proposed ASTODE model obtains a similar or higher prediction accuracy. This paper also analyzes the interpretability of our proposed ASTODE model from a visualization perspective, demonstrating that the proposed ASTODE model balances the prediction accuracy and interpretability to some extent.

**Key words:** PM<sub>2.5</sub> concentration prediction; air pollution; energy conservation and emission reduction; spatiotemporal prediction; attention mechanism; Neural Ordinary Differential Equation; Spatiotemporal Ordinary Differential Equation; model interpretability

\*Corresponding author: ZHANG Hengcai, E-mail: zhanghc@lreis.ac.cn

**摘要** 可解释的准确预测PM<sub>2.5</sub>浓度变化可以有助于人类规避暴露风险,对人类健康风险评估和政策实施具有重要意义。目前已有PM<sub>2.5</sub>浓度预测模型过多专注于提升模型预测精度,但忽略了模型的可解释性,造成模型可复用性和可信任度较差。鉴于此,本文提出了一种兼顾模型预测精度与模型可解释性的注意力时空常微分方程模型(Attentional SpatioTemporal Ordinary Differential Equation, ASTODE)用于PM<sub>2.5</sub>浓度预测任务。具体而言,本文将神经常微分方程集成至PM<sub>2.5</sub>浓度预测任务中,以提升预测模型的可解释性。此外,针对传统神经常微分方程难以挖掘PM<sub>2.5</sub>浓度数据中空间依赖关系的挑战,本文提出了一种新颖时空导数网络将传统神经常微分方程扩展到了时空常微分方程。针对传统神经常微分方程难以挖掘PM<sub>2.5</sub>浓度数据中长期依赖关系的挑战,本文设计了一种时空注意力机制去融合多个时间节点的隐藏状态。本文采用真实的PM<sub>2.5</sub>浓度数据集对提出的ASTODE模型进行了验证。实验结果表明,ASTODE模型不仅在预测精度上优于或逼近于存在的6个基线方法,并且在可视化的视角下具有良好的可解释性。

**关键词** PM<sub>2.5</sub>浓度预测;空气污染;节能减排;时空预测;注意力机制;神经常微分方程;时空常微分方程;模型可解释性

## 1 引言

PM<sub>2.5</sub>作为空气污染物的主要成分之一,严重影响人类身体健康,如呼吸系统受损、癌症与心血管疾病的发病率增加等<sup>[1-2]</sup>。如何降低人类在PM<sub>2.5</sub>中暴露程度是当下可持续城市和社会发展的重中之重<sup>[3-5]</sup>,其中PM<sub>2.5</sub>浓度预测技术就是降低人类暴露风险的重要手段之一。准确可靠的预测PM<sub>2.5</sub>浓度变化可以有助于人类规避暴露风险,辅助环保相关部门预警并制定决策<sup>[6-7]</sup>。

PM<sub>2.5</sub>浓度预测属于典型的时空预测任务,可分为知识驱动预测模型与数据驱动预测模型<sup>[8-10]</sup>。相较于知识驱动预测模型,数据驱动预测模型通过训练复杂的机器学习或深度学习模型,建立输入数据与输出数据之间的非线性映射关系,进而提升时空

预测精度<sup>[11-13]</sup>,例如时空k近邻模型(Spatiotemporal k-nearest neighbor, ST-KNN)<sup>[14]</sup>、时间图卷积网络模型(Temporal Graph Convolutional Network, T-GCN)<sup>[15]</sup>、贝叶斯矩阵分解模型(Bayesian Temporal Matrix Factorization, BTMF)<sup>[16]</sup>、通用动态图卷积网络(Generic Dynamic Graph Convolutional Network, GDGCN)<sup>[17]</sup>。尽管数据驱动的预测模型已经在时空预测任务中取得了较优的预测精度,但仍面临挑战。具体而言,数据驱动预测模型专注于预测精度的提升,模型可解释性相对较差<sup>[18]</sup>。可解释性PM<sub>2.5</sub>浓度预测一直是学界与业界关注的重点,提高预测精度同时增强预测模型的可解释性、透明度是减少黑盒场景,提高模型信任度与复用性的重中之重<sup>[19-20]</sup>。但大多数时空预测模型却难以兼顾预测精度及可解释之间的均衡<sup>[21-22]</sup>。

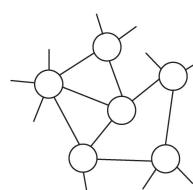
近年来,神经常微分方程(Neural Ordinary Differential Equation, NODE)兴起为可解释的PM<sub>2.5</sub>浓度预测研究提供了一种新的解决思路<sup>[23-25]</sup>。NODE模型利用神经网络参数化导数网络构建深度学习和常微分方程之间关系,进而提高PM<sub>2.5</sub>浓度预测可解释性<sup>[24]</sup>。目前已经有学者提出了NODE变种模型用于时空预测任务,例如循环神经网络常微分方程(Recurrent Neural Network Ordinary Differential Equations, ODE-RNNs)<sup>[26]</sup>。尽管传统NODE模型及其变种模型兼顾了模型预测精度及可解释之间的均衡,但在PM<sub>2.5</sub>浓度预测任务中仍面临着挑战。究其原因在于,传统NODE模型及其变种模型在提升模型可解释性的同时,牺牲了模型预测精度<sup>[27]</sup>。上述模型是一个单纯的时间序列模型,难以挖掘PM<sub>2.5</sub>浓度数据中空间依赖关系,且过度依赖于常微分方程初值,难以挖掘PM<sub>2.5</sub>浓度数据中长期依赖关系。鉴于此,本文提出了一种新颖的兼顾预测精度和可解释性的PM<sub>2.5</sub>浓度预测模型,即注意力时空常微分方程(Attentional SpatioTemporal Ordinary Differential Equation, ASTODE)。首先,通过设计了一种新颖时空导数网络将传统NODE扩展到了时空常微分方程,进而挖掘PM<sub>2.5</sub>浓度数据中的时空关系,接着,构建了一种可解释的数据结构(即注意力机制)辅助时空常微分方程捕捉PM<sub>2.5</sub>浓度数据中的长期依赖关系,从而实现在保证模型可解释前提下去提升模型预测精度。

## 2 问题与定义

本小节给出ASTODE模型所需要的相关定义及PM<sub>2.5</sub>浓度预测问题的数学描述。



(a) 研究区域



(b) 图结构

● 空气质量监测站点 ○ 图节点  $\boxed{x_i^t}$  监测站点 $v_i$ 在时间窗口 $t$ 采集的PM<sub>2.5</sub>浓度数据

**定义1(图):**如图1所示,图 $G=<V,E,A>$ 表示由空气质量站点抽象为的图结构,其中 $V=\{v_i\}_{i=1}^n$ 表示图 $G$ 中的 $n$ 个空气质量监测站点; $E=\{e_{ij}\}$ 表示节点 $v_i$ 和节点 $v_j$ 之间的拓扑连接关系。为简单,空气质量监测站点之间的拓扑连接关系用矩阵 $A \in R^{n \times n}$ 表示,其中 $A_{ij}=1$ 表示空气质量监测站点 $v_i$ 和 $v_j$ 直接相连, $A_{ij}=0$ 表示空气质量监测站点 $v_i$ 和 $v_j$ 不相连。在本研究中,监测站点之间的距离矩阵用于构建拓扑连接矩阵 $A$ 。具体而言,本文利用距离矩阵将5个最近的监测站点视为目标监测站点的拓扑邻居。

**定义2(PM<sub>2.5</sub>浓度矩阵):**PM<sub>2.5</sub>浓度矩阵可由 $X \in R^{n \times T}$ 表示,其中 $x_i^t \in R^{1 \times 1}$ 表示空气质量监测站点 $v_i$ 在时间窗口 $t$ 内采集到的PM<sub>2.5</sub>浓度, $x^t = \{x_{ij}^t\}_{i=1}^n \in R^{n \times 1}$ 表示所有空气质量监测站点在时间窗口 $t$ 内采集的PM<sub>2.5</sub>浓度, $n$ 表示空气质量监测站点的个数, $T$ 表示所有时间窗口的个数。

本文的研究目标旨在建立一个函数模型 $F(\cdot)$ ,该模型可以从PM<sub>2.5</sub>浓度矩阵 $X$ 中挖掘时空数据的模式,从而可解释的预测未来的PM<sub>2.5</sub>浓度信息。给定一个PM<sub>2.5</sub>浓度矩阵 $X$ ,其建模过程如式(1)所示。

$$\hat{x}^{t+1} = F(X_{t-p+1}^t, G; W) \quad (1)$$

式中: $F(\cdot)$ 表示本文提出ASTODE模型; $G$ 表示由空气质量站点抽象为的图结构; $X_{t-p+1}^t = \{x^{\tau}\}_{\tau=t-p+1}^t \in R^{n \times p}$ 表示ASTODE模型需要输入的历史PM<sub>2.5</sub>浓度数据,其中 $p$ 表示时间依赖步长; $\hat{x}^{t+1}$ 表示将来(预测)的PM<sub>2.5</sub>浓度数据; $W$ 表示ASTODE模型中可学习的参数。

|       | 1 | ... | $t$     | ... | $T$ |
|-------|---|-----|---------|-----|-----|
| $v_n$ |   |     |         |     |     |
| $v_i$ |   |     | $x_i^t$ |     |     |
| $v_1$ |   |     |         |     |     |

(c) PM<sub>2.5</sub>浓度矩阵

图1 ASTODE模型所需要的相关定义

Fig. 1 Relevant definitions required by the ASTODE model

### 3 注意力时空常微分方程

图2为ASTODE模型整体流程,由2个核心组件组成:时空常微分方程模块和时空注意力模块。在时空常微分方程模块中,将每个时刻的隐藏状态定义为常微分方程的解,依据定义时空导数网络的可解释的求解每一个时刻的隐藏状态(3.1节)。在时空注意力模块中,采用分别空间注意力和时间注意力去融合多个历史时刻的隐藏状态,从而捕捉时空数据中的长期依赖关系(3.2节)。最后,通过最小化模型预测结果与观测真值之间的损失去优化ASTODE模型中的参数(3.3节)。

#### 3.1 时空常微分方程模块

本小节首先介绍了传统神经常微分方程的数学原理及其可解释性,然后在传统神经常微分方程的基础上提出了时空常微分方程模块。

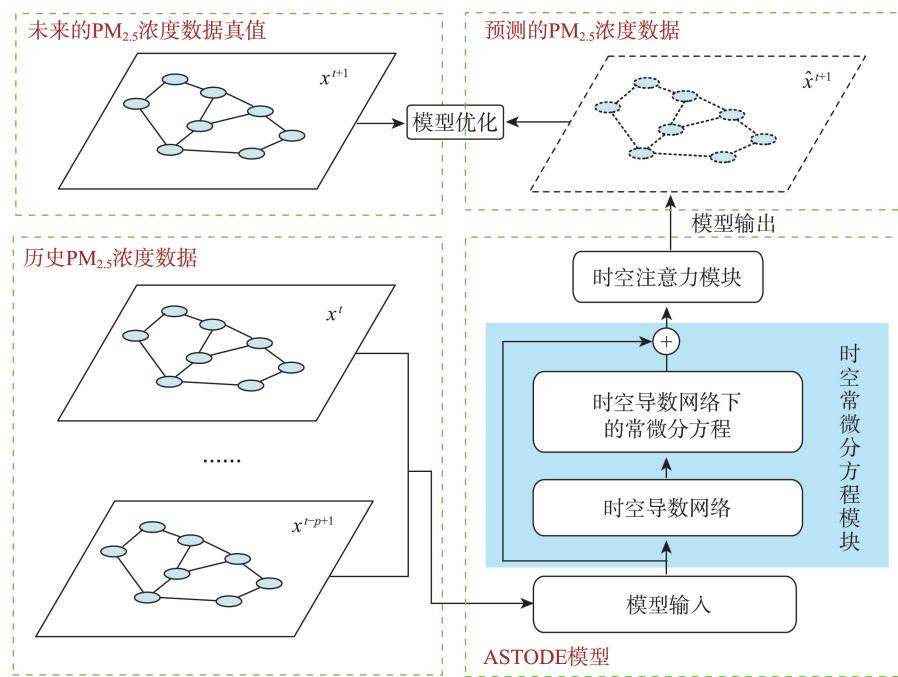
传统NODE<sup>[23]</sup>是一类连续时间域的时间序列预测模型,其针对单个监测站点监测的时间序列建模,其将单个监测站点每个时刻的观测值视为常微分方程的解,依据导数网络迭代的求解预测值。假设时间序列  $\mathbf{x}_i = \{x_i^t\}_{t=1}^T \in R^{T \times 1}$  为节点  $v_i$  在连续时间域

$x(t)$  的离散采样,监测站点在特定时间下的预测值如式(2)所示。

$$x_i(t) = x_i(0) + \int_0^t \frac{dx_i(\tau)}{d\tau} d\tau = x_i(0) + \int_0^t g(x_i(\tau), \tau) d\tau \quad (2)$$

式中:  $x_i(t) \in R^{1 \times 1}$  表示监测站点  $v_i$  在  $t$  时刻的预测值;  $x_i(0)$  表示神经常微分方程的初值,即监测站点  $v_i$  在 0 时刻的监测值;  $g(x_i(t), t) = \frac{dx_i(t)}{dt}$  表示导数网络,其通过数值ODE求解器求解<sup>[23]</sup>。从模型可解释的角度来看,  $g(x_i(t), t)$  表示连续时间域中函数  $x_i(t)$  沿时间  $t$  的导数,即模型预测值  $x_i(t)$  在模型初值  $x_i(t_0)$  的基础上,依据导数网络  $g(x_i(t), t)$  多次上坡和下坡获得。由于NODE的预测过程具有明确的数学机理,所以本文认为NODE是具备一定可解释的预测模型。由式(2)可知,传统NODE虽具备一定的可解释性,但仍存在不足:①传统NODE的导数网络  $g(x_i(t), t)$  仅依赖时间因素,没有考虑多监测站点的情形;②传统NODE在求解过程中仅依赖初值  $x_i(0)$ ,使该模型难以挖掘PM<sub>2.5</sub>数据中长期依赖关系。

本小节主要针对第一点不足,提出了时空导数网络将传统NODE扩展到了时空常微分方程。此外,为了提升ASTODE模型的非线性拟合能力,本



注:  $x^t \dots x^{t-p+1}$  表示历史PM<sub>2.5</sub>浓度数据;  $x^{t+1}$  表示未来的PM<sub>2.5</sub>浓度数据真值;  $\hat{x}^{t+1}$  表示预测的PM<sub>2.5</sub>浓度数据。

图2 ASTODE模型的整体示意

Fig. 2 Overall schematic of the ASTODE model

文将每个时刻的隐藏状态视为常微分方程的解,而非将每个时刻观测值视为常微分方程的解。时空导数网络数学定义如式(3)所示。

$$g(\mathbf{H}(t), t, \{\mathbf{i}\}_{i=1}^n) \quad (3)$$

式中:  $\mathbf{H}(t)=\{\mathbf{h}_i(t)\}_{i=1}^n \in R^{n \times d_h}$  表示  $n$  个监测站点在时刻  $t$  的隐藏状态, 其中  $\mathbf{h}_i(t) \in R^{1 \times d_h}$  表示第  $i$  个监测站点在时刻  $t$  的隐藏状态,  $\{\mathbf{i}\}_{i=1}^n$  表示不同监测站点的空间位置信息, 其在时空导数网络内部进一步被编码。由于显式的建模了空间信息, 所以常微分方程扩展到了时空常微分方程。

时空常微分方程模块理论上即可根据式(2)迭代的求解每一个时刻的隐藏状态, 式(2)是纯粹的迭代模型, 迭代模型在优化过程中易造成梯度消失\爆炸现象, 从而导致模型优化缓慢, 甚至是无法优化。为了缓解优化过程中的梯度消失\爆炸现象, 在时空常微分方程模块中引入了门控机制<sup>[28]</sup>和跳跃连接<sup>[29]</sup>加快模型的优化效率(图3)。以PM<sub>2.5</sub>浓度矩阵  $\mathbf{X}'_{t-p+1}=\{\mathbf{x}'^{t-p+1}, \mathbf{x}'^{t-p+2}, \dots, \mathbf{x}'\}$  为例, 每个时刻隐藏状态的迭代求解过程如公式(4)所示, 观测值与隐藏状态的残差连接如式(5)所示。

$$\begin{cases} \mathbf{Z}'^{t-1} = \sigma(\mathbf{W}_z[\mathbf{H}'^{t-1} \parallel \mathbf{x}']) \\ \mathbf{I}'^{t-1} = \sigma(\mathbf{W}_i[\mathbf{H}'^{t-1} \parallel \mathbf{x}']) \\ \tilde{\mathbf{C}}'^{t-1} = \tanh(\mathbf{W}_r[\mathbf{H}'^{t-1} \parallel \mathbf{x}']) \\ \mathbf{C}'^{t-1} = \mathbf{C}'^{t-2} \odot \mathbf{I}'^{t-1} + \mathbf{Z}'^{t-1} \odot \tilde{\mathbf{C}}'^{t-1} \\ \mathbf{L}'^{t-1} = \sigma(\mathbf{W}_c[\mathbf{H}'^{t-1} \parallel \mathbf{x}']) \\ \tilde{\mathbf{H}}'^{t-1} = \mathbf{L}'^{t-1} \odot \tanh(\mathbf{C}'^{t-1}) \\ \mathbf{H}' = \tilde{\mathbf{H}}'^{t-1} + \int_{t-1}^t g(\mathbf{H}, \tau, \{\mathbf{i}\}_{i=1}^n) d\tau \end{cases} \quad (4)$$

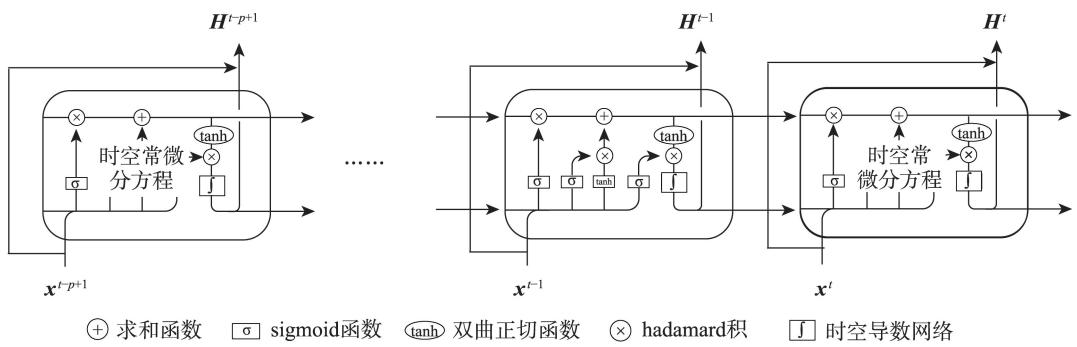
$$\mathbf{H}=[\mathbf{H}'^{t-p+1} \parallel \dots \parallel \mathbf{H}'] + Conv(\mathbf{X}'_{t-p+1}) \quad (5)$$

式中:  $\mathbf{H} \in R^{n \times d_h \times p}$  表示跳跃连接后的  $p$  个时刻的隐藏状态;  $\mathbf{H}' \in R^{n \times d_h}$  表示所有监测站点在  $t$  时刻迭代求解的隐藏状态,  $d_h$  表示隐藏状态的维度;  $\mathbf{x}' \in R^{n \times 1}$  表示所有监测站点在  $t$  时刻的观测数据;  $\mathbf{Z}'^{t-1}$ 、 $\mathbf{I}'^{t-1}$ 、 $\mathbf{L}'^{t-1}$  表示门控, 即输出门、遗忘门、输入门;  $\mathbf{C}'^{t-1}$ 、 $\tilde{\mathbf{C}}'^{t-1}$ 、 $\tilde{\mathbf{H}}'^{t-1}$  表示中间变量;  $\sigma$  表示 sigmoid 激活函数;  $\tanh$  表示 tanh 激活函数;  $\odot$  表示 hadamard 积;  $[\cdot \parallel \cdot]$  表示 concatenate 函数;  $g(\mathbf{H}, \tau, \{\mathbf{i}\}_{i=1}^n)$  表示时空导数网络;  $Conv$  表示用于跳跃连接的卷积运算;  $\mathbf{W}$  表示时空常微分模块中可学习的权重。

### 3.2 时空注意力模块

本小节主要针对传统 NODE 难以挖掘 PM<sub>2.5</sub> 数据中长期依赖关系的挑战, 提出一种时空注意力模块去融合时空常微分方程模块的输出结果。之所以采用注意力机制, 是因为注意力机制是一种可解释的数据结构。以空间维度的注意力机制为例, 注意力机制可视为反距离权重函数在数据驱动下的一种扩展(图4)。具体而言, 反距离权重中的“权重”采用距离计算, 而注意力机制中的“权重”通过神经网络求得, 这使得注意力机制在具备可解释的前提下拥有更强的非线性拟合能力, 同时也是本文采用注意力机制融合时空常微分方程模块输出结果的原因。

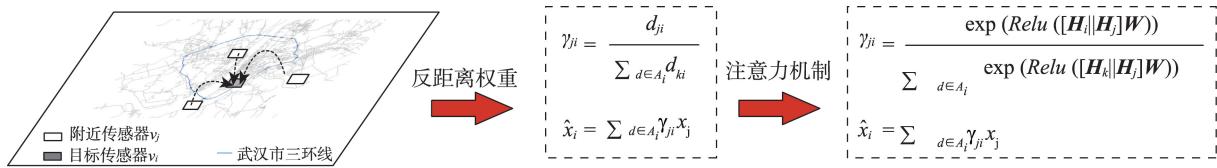
时空注意力模块由空间注意力和时间注意力组成, 其中空间注意力用于在空间维度融合  $n$  个监测站点的隐藏状态, 时间注意力用于在时间维度融合  $p$  个时间窗口的隐藏状态。如图5所示, 在空间注意力中, 三维张量  $\mathbf{H} \in R^{n \times d_h \times p}$  的维度将变化为二维矩阵  $\mathbf{H}^s \in R^{n \times (d_h \times p)}$ ; 在时间注意力中, 三维张量  $\mathbf{H} \in R^{n \times d_h \times p}$



注:  $\mathbf{H}^{t-1}$  表示所有监测站点在时间窗口  $t-1$  的隐藏状态;  $\mathbf{x}^{t-1}$  表示所有监测站点在时间窗口  $t-1$  采集的PM<sub>2.5</sub>浓度。

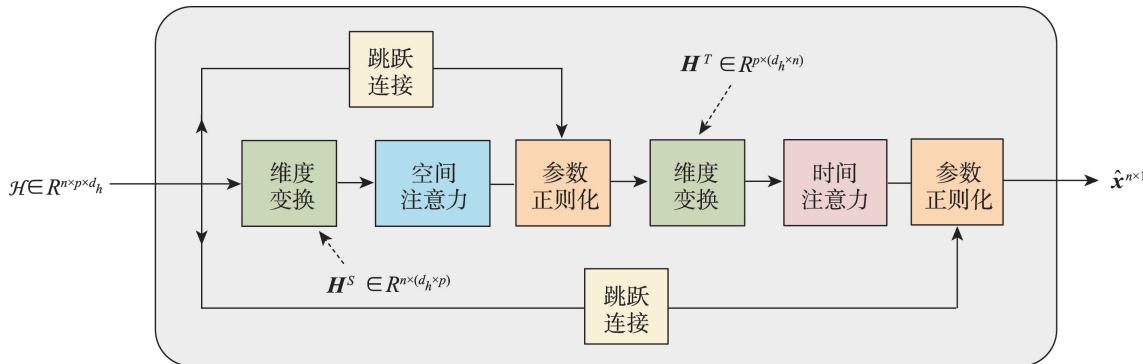
图3 时空常微分方程模块的前向传播过程

Fig. 3 Forward propagation of spatiotemporal ordinary differential equation



注:  $d_{ji}$  表示传感器  $v_i$  和  $v_j$  之间的距离;  $\gamma_{ji}$  为传感器  $v_i$  对  $v_j$  的影响权重;  $\hat{x}_i$  表示传感器  $v_i$  的预测值;  $x_j$  表示传感器  $v_j$  的观测值;  $\mathbf{H}_i$  表示传感器  $v_i$  的隐藏状态;  $\mathbf{W}$  表示注意力机制中的可学习参数。

图4 注意力机制的可解释性  
Fig. 4 Interpretability of attention mechanisms



注:  $\mathcal{H}$  表示模块的输入张量(三维);  $\hat{\mathbf{x}}^{n \times 1}$  表示模块的输出矩阵(二维);  $\mathbf{H}^S$  表示空间注意力的输入矩阵(二维);  $\mathbf{H}^T$  表示时间注意力的输入矩阵(二维)。

图5 时空注意力模块的前向传播过程  
Fig. 5 Forward propagation of spatiotemporal attention

的维度将变化为二维矩阵  $\mathbf{H}^T \in R^{p \times (d_h \times n)}$ 。为加快模型的优化效率,本文也将跳跃连接集成到时空注意力模块去加快模型的优化效率。具体而言,空间注意力和时间注意力的融合过程如式(6)、式(7)所示。

$$\begin{cases} \mathbf{O}_i^S = \sum_{j \in A_i} \gamma_{ji} \mathbf{H}_j^S \mathbf{W}_v \\ \gamma_{ji}^S = \frac{\exp(\text{Relu}([\mathbf{H}_i^S \| \mathbf{H}_j^S] \mathbf{W}_q))}{\sum_{k \in A_i} \exp(\text{Relu}([\mathbf{H}_i^S \| \mathbf{H}_k^S] \mathbf{W}_q))} \end{cases} \quad (6)$$

$$\begin{cases} \mathbf{O}^{T:t} = \sum_{\tau=t-p+1}^t \gamma^{T:\tau \rightarrow t} \mathbf{H}^{T:\tau} \mathbf{W}_v \\ \gamma^{T:\tau \rightarrow t} = \frac{\exp(\text{Relu}([\mathbf{H}^{T:\tau} \| \mathbf{H}^{T:t}] \mathbf{W}_q))}{\sum_{k=t-p+1}^t \exp(\text{Relu}([\mathbf{H}^{T:k} \| \mathbf{H}^{T:t}] \mathbf{W}_q))} \end{cases} \quad (7)$$

式中:  $\mathbf{O}_i^S$  表示监测站点  $v_i$  的隐藏状态经空间注意力的结果;  $\mathbf{O}^{T:t}$  表示第  $t$  个时间窗口的隐藏状态经时间注意力的结果;  $\mathbf{H}_i^S$  表示监测站点  $v_i$  在多个时间窗口隐藏状态的拼接;  $\mathbf{H}^{T:t}$  表示第  $t$  个时间窗口所有

监测站点的拼接;  $A_i$  表示监测站点  $v_i$  的拓扑邻接关系;  $\odot$  表示 hadamard 积;  $[\cdot \| \cdot]$  表示 concatenate 函数;  $\exp$  表示指数函数;  $\text{Relu}$  表示激活函数;  $\mathbf{W}$  表示模块中可学习的权重。

### 3.3 模型优化

本研究通过前  $p$  个时间窗口的 PM<sub>2.5</sub> 浓度数据  $\{\mathbf{x}^{\tau}\}_{\tau=t-p+1}^t$  去预测将来的 PM<sub>2.5</sub> 浓度数据  $\hat{\mathbf{x}}^{t+1}$ 。考虑到 ASTODE 模型是一种端到端的监督学习模型,因此,本文仅需要优化预测值  $\hat{\mathbf{x}}^{t+1}$  与真值  $\mathbf{x}^{t+1}$  之间的均方误差即可获得最终的预测模型。ASTODE 模型损失函数如式(8)所示。

$$L(\mathbf{W}) = \min_{\mathbf{W}} (\mathbf{x}^{t+1} - \hat{\mathbf{x}}^{t+1})^2 \quad (8)$$

式中:  $\mathbf{x}^{t+1} \in R^{n \times 1}$  表示 ASTODE 模型在第  $t+1$  个时间窗口的 PM<sub>2.5</sub> 浓度数据真值;  $\hat{\mathbf{x}}^{t+1} \in R^{n \times 1}$  表示 ASTODE 模型在第  $t+1$  个时间窗口的 PM<sub>2.5</sub> 浓度数据预测值;  $\odot$  表示 Hadamard 积,即逐个元素相乘;  $\mathbf{W}$  表示模型中优化的参数。

## 4 实验及分析

### 4.1 实验数据及定量评价指标

本研究采用的PM<sub>2.5</sub>浓度数据集来源于天津市的27个空气质量监测站点(图6)<sup>[30]</sup>,时间跨度为2015年2月1日—2015年4月30日,采样间隔为60 min。表1展示了单个空气质量监测站点采集的PM<sub>2.5</sub>浓度数据实例,每条记录包含了加密后的监测站点唯一标识、监测的时间窗口以及时间窗口内的PM<sub>2.5</sub>浓度信息。在数据预处理阶段,本文采用贝叶斯张量分解模型<sup>[16]</sup>补全原始数据自然缺失值,以免自然缺失值影响ASTODE模型的预测精度。接着将预处理后的PM<sub>2.5</sub>浓度数据集分为训练数据集和测试数据集,其中训练数据集占80%,测试数据集占20%。

本研究采用平均绝对误差(Mean Absolute Error,

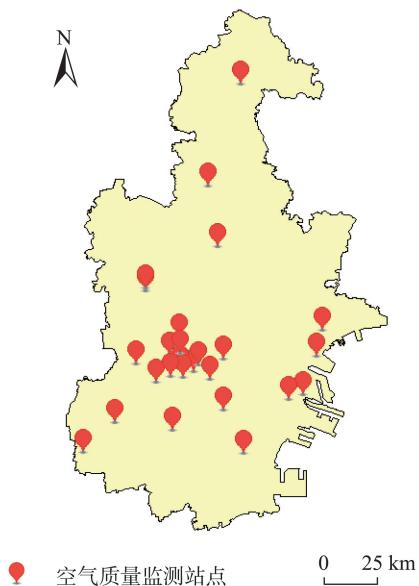


图6 天津市空气质量监测站点空间分布

Fig. 6 Spatial distribution of air quality monitoring stations in Tianjin

表1 PM<sub>2.5</sub>浓度数据实例

Tab. 1 Examples of PM<sub>2.5</sub> concentration data

| 监测站点<br>唯一标识 | 时间窗口                              | PM <sub>2.5</sub> 浓度 |
|--------------|-----------------------------------|----------------------|
| b5a6da1e6w   | 2015-02-01 00:00—2015-02-01 01:00 | 71                   |
| b5a6da1e6w   | 2015-02-01 01:00—2015-02-01 02:00 | 66                   |
| b5a6da1e6w   | 2015-02-01 02:00—2015-02-01 03:00 | 58                   |
| .....        | .....                             | .....                |
| b5a6da1e6w   | 2015-04-30 22:00—2015-04-30 23:00 | 112                  |

MAE)、均方根误差(Root Mean Square Error, RMSE)及平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)作为定量指标来验证所提出模型的预测精度。MAE、RMSE、MAPE的计算方法如式(9)、式(11)所示。

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i^{t+1} - \hat{x}_i^{t+1}| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i^{t+1} - \hat{x}_i^{t+1})^2} \quad (10)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{x_i^{t+1} - \hat{x}_i^{t+1}}{x_i^{t+1}} \right| \quad (11)$$

式中: $x_i^{t+1}$ 表示监测节点 $v_i$ 在 $t+1$ 时间窗口内的真实PM<sub>2.5</sub>浓度数据,即真值; $\hat{x}_i^{t+1}$ 表示监测节点 $v_i$ 在 $t+1$ 时间窗口内预测的PM<sub>2.5</sub>浓度数据; $n$ 表示研究区域中监测站点的总个数。

### 4.2 模型预测精度

在本小节中,图7采用折线图和地图定性描述了ASTODE模型的预测性能。图7(a)从时间维度描述了预测值和真值之间的差异。整体而言,在时间维度上,大多数时刻观测真值与模型预测值的残差较小,少量时刻观测真值与模型预测值的残差较大(蓝色区域)。ASTODE模型之所以在蓝色区域内的残差较大,主要原因是蓝色区域内PM<sub>2.5</sub>浓度的走势在短时间内发生了突变。短时间内的PM<sub>2.5</sub>浓度的走势发生突变可能会引入观测误差,导致模型预测准确的难度加大。此外,图7(b)从空间维度描述了预测值和真值之间的差异。在空间维度上,大多数监测站点观测真值与模型预测值的残差较小,少量监测站点观测真值与模型预测值的残差较大(蓝色区域)。同理,ASTODE模型之所以在蓝色区域内的残差较大,主要是因为蓝色区域中的监测站点PM<sub>2.5</sub>浓度波动较大(预测误差与数据波动存在明显的正相关),这增加了模型预测准确的难度。

### 4.3 基准方法对比分析

本文采用的基线方法大致可以分为2类。第一类是纯黑盒的数据驱动模型,包括ST-KNN模型<sup>[14]</sup>、T-GCN模型<sup>[15]</sup>、BTMF模型<sup>[16]</sup>、GDGCN模型<sup>[17]</sup>。第二类是具备一定可解释性的常微分方程及其变种模型,包括传统NODE模型(Vanilla-NODE)<sup>[23]</sup>和ODE-RNNs模型<sup>[26]</sup>。

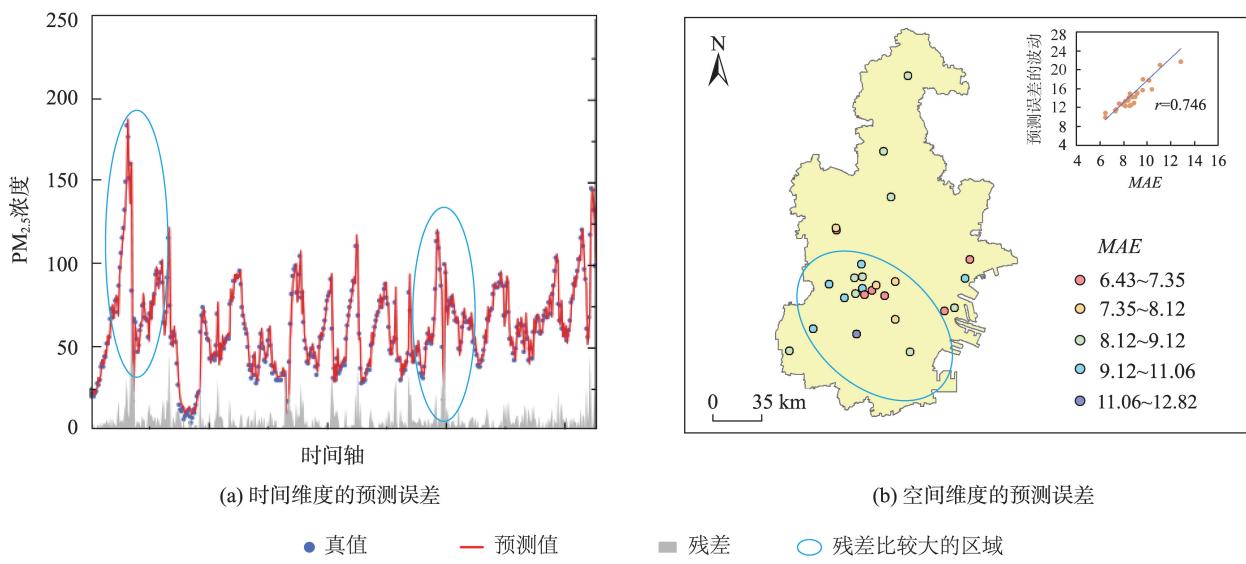


图7 ASTODE模型的定性预测结果  
Fig. 7 Qualitative prediction results of ASTODE model

表2展示了ASTODE模型和基准方法在真实PM<sub>2.5</sub>浓度数据集中的对比结果。结果表明,第一类模型的预测精度在近年来获得了较大的提升,且优于第二类模型的预测精度。造成上述结果原因除了基于NODE的预测模型提升模型可解释性的同时,牺牲了模型的预测精度。具体而言,基于NODE的预测模型的导数网络多数仅依赖于时间因素,并没有将空间位置信息显式的建模到导数网络的输出中。其次,基于NODE的预测模型依赖于ODE的初值,难以发现时空数据中的长时间依赖关系。相较于基准实验,ASTODE模型解决了上述的2点不足,兼顾模型可解释的同时,优于或者逼近于现有基线方法的预测精度。尤其是ASTODE模型已经逼近于GDGCN模型的预测精度,这证明了在不考虑模型可解释性的情况下,

表2 ASTODE与基准方法的比较结果  
Tab. 2 Comparison results of prediction performance between ASTODE and baselines

| 模型           | MAE   | RMSE  | MAPE/% |
|--------------|-------|-------|--------|
| ST-KNN       | 14.35 | 22.12 | 32.02  |
| T-GCN        | 10.20 | 17.97 | 25.14  |
| BTMF         | 8.98  | 15.42 | 20.74  |
| GDGCN        | 8.56  | 15.11 | 19.47  |
| Vanilla-NODE | 13.34 | 19.52 | 35.06  |
| ODE-RNNs     | 9.67  | 15.99 | 23.22  |
| ASTODE       | 8.63  | 15.14 | 19.54  |

ASTODE模型已经具备了令人可接受的预测结果。

#### 4.4 模型可解释性分析

本研究从2个方面解释ASTODE模型:①解释时空常微分方程模块学习到的导数值;②解释时空注意力模块学习到时空关系。

考虑到隐藏状态导数值难以理解,本文将每个时刻观测值视为常微分方程的解,单独训练PM<sub>2.5</sub>浓度预测模型,进而可视化了观测值的导数值,结果如图8(a)所示。结果表明10:45之前和19:45之后是PM<sub>2.5</sub>上升的主要时间段,10:45分—19:45是PM<sub>2.5</sub>下降的主要时间段,这恰好与可视化导数值相对应,从而证明了时空常微分方程模块具备一定的可解释性。本研究进一步可视化了时空注意力模块学习到时空关系,结果如图8(b)和图8(c)所示。结果表明,在时间维度,t时刻的预测值受到历史3个时刻的影响,且历史时刻距离目标时刻越近,对预测值的影响权重越大。在空间维度,邻域监测站点对目标监测站点的影响权重存在明显差异,且距离目标监测站点越近的邻域监测站点对目标监测站点的影响权重越大。时空注意力模块便是在学习时间维度和空间维度的影响权重,进而用于PM<sub>2.5</sub>浓度预测任务,这也证明了时空注意力模块具备一定的可解释性。

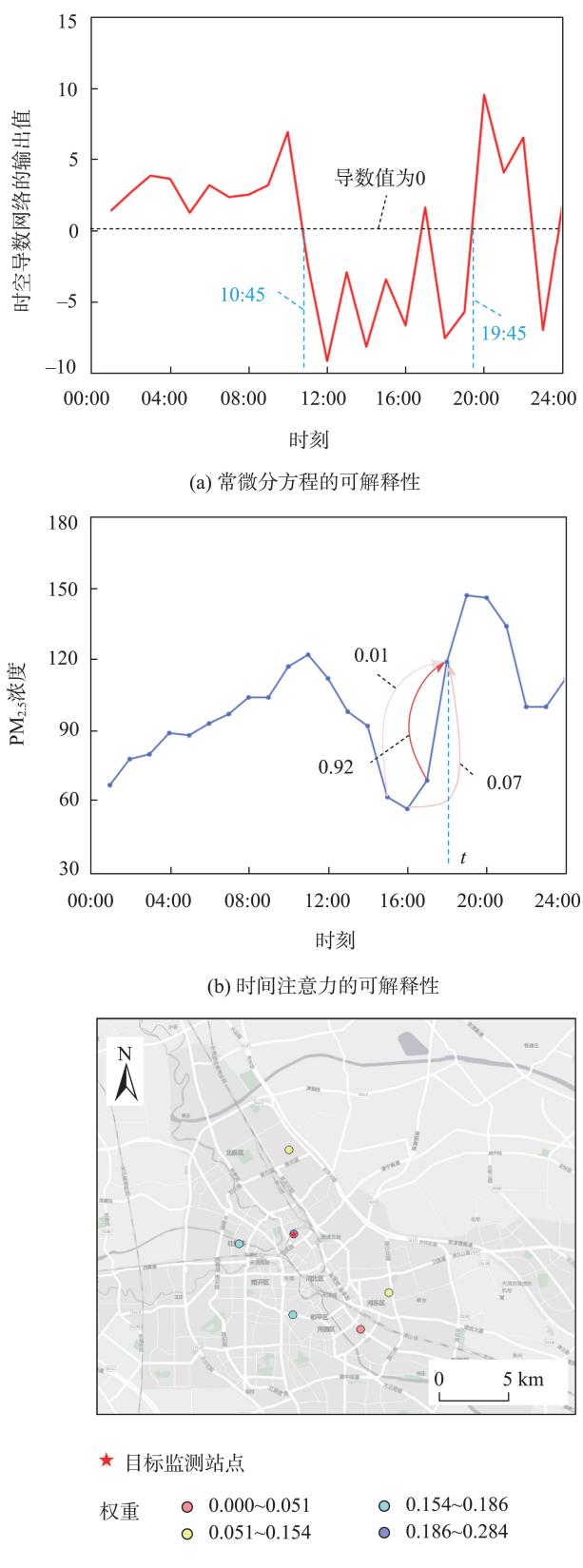


图8 ASTODE模型的可解释性

Fig. 8 Interpretability of the ASTODE model

## 5 结论与展望

可解释的准确预测PM<sub>2.5</sub>浓度变化对人类健康风险评估和政策实施具有重要意义。然而,现有的大多数PM<sub>2.5</sub>浓度预测模型专注于提升模型的预测精度,但忽略了模型的可解释性。鉴于此,本文提出了一种新颖的兼顾预测精度及可解释的ASTODE模型。具体而言,本文将神经常微分方程集成到时空预测任务中,以提升模型的可解释性。针对传统神经常微分方程面向的2个挑战,本文提出了时空导数网络和时空注意力机制去提升模型捕捉时空关系的能力及捕捉长期依赖关系的能力。

在实验部分,本文使用真实的PM<sub>2.5</sub>浓度数据集验证了ASTODE模型的预测性能。首先,本文从定性的角度分析了ASTODE模型在时间维度和空间维度的预测误差;其次,本文与现有的6个基线方法(ST-KNN、T-GCN、BTMF、GDGCN、Vanilla-NODE和ODE-RNNs)进行了比较。实验结果表明,ASTODE模型在预测精度上优于或者等于存在的先进基线方法。最后,本文从可视化的角度分析了ASTODE模型的可解释性,证明了ASTODE模型在一定程度上兼顾了模型的预测精度及可解释。

本研究的尚有如下的局限性:①ASTODE模型基于ODE完成时空预测任务,ODE前向迭代和反向传播是一个十分耗时的过程,进而导致ASTODE模型的训练更加困难;②本文仅验证了ASTODE模型的单步预测性能,并未验证ASTODE模型多步预测能力。针对上述问题,下一步工作重点将从两方面展开。首先,进一步优化ASTODE模型的前向迭代和反向传播的迭代过程,提升ASTODE模型的训练效率。其次,进一步验证ASTODE模型多步预测能力。

**致谢:**本文的数值计算得到了武汉大学超级计算中心的计算支持和帮助。

### 参考文献(References):

- [1] Feng S L, Gao D, Liao F, et al. The health effects of ambient PM<sub>2.5</sub> and potential mechanisms[J]. Ecotoxicology and Environmental Safety, 2016,128:67-74. DOI:10.1016/j.ecoenv.2016.01.030
- [2] Kim K H, Kabir E, Kabir S. A review on the human

- health impact of airborne particulate matter[J]. Environment International, 2015,74:136-143. DOI:10.1016/j.envnt.2014.10.005
- [3] Yang K, Teng M F, Luo Y, et al. Human activities and the natural environment have induced changes in the PM<sub>2.5</sub> concentrations in Yunnan Province, China, over the past 19 years[J]. Environmental Pollution, 2020,265:114878. DOI:10.1016/j.envpol.2020.114878
- [4] 冯子钰,施润和.中国近地面PM<sub>2.5</sub>浓度与排放的时空分布及其关联分析[J].地球信息科学学报,2021,23(7):1221-1230. [ Feng Z Y, Shi R H. Spatio-temporal features and the association of ground-level PM<sub>2.5</sub> concentration and its emission in China[J]. Journal of Geo-information Science, 2021,23(7):1221-1230. ] DOI:10.12082/dqxxkx.2021.200367
- [5] 杜震洪,吴森森,王中一,等.基于地理神经网络加权回归的中国PM<sub>2.5</sub>浓度空间分布估算方法[J].地球信息科学学报,2020,22(1):122-135. [ Du Z H, Wu S S, Wang Z Y, et al. Estimating ground-level PM<sub>2.5</sub> concentrations across China using geographically neural network weighted regression[J]. Journal of Geo-information Science, 2020,22(1):122-135. ] DOI:10.12082/dqxxkx.2020.190533
- [6] Teng M F, Li S W, Song G, et al. Including the feature of appropriate adjacent sites improves the PM<sub>2.5</sub> concentration prediction with long short-term memory neural network model[J]. Sustainable Cities and Society, 2022,76:103427. DOI:10.1016/j.scs.2021.103427
- [7] Zheng Q H, Tian X Y, Yu Z G, et al. Application of wavelet-packet transform driven deep learning method in PM<sub>2.5</sub> concentration prediction: A case study of Qingdao, China [J]. Sustainable Cities and Society, 2023,92:104486. DOI:10.1016/j.scs.2023.104486
- [8] Chen Y F, Guizani M, Zhang Y, et al. When traffic flow prediction and wireless big data analytics meet[J]. IEEE Network, 2019,33(3):161-167. DOI:10.1109/MNET.2018.1800134
- [9] 毛文婧,王卫林,焦利民,等.基于深度学习的中国连续空间覆盖PM<sub>2.5</sub>浓度预报[J].测绘学报,2022,51(3):361-372. [ Mao W J, Wang W L, Jiao L M, et al. Continuous spatial coverage PM<sub>2.5</sub> concentration forecast in China based on deep learning[J]. Acta Geodaetica et Cartographica Sinica, 2022,51(3):361-372. ] DOI:10.11947/j.AGCS.2022.2020038
- [10] 赵琨,罗力,杨凤芸.利用北斗 CORS 反演大气可降水量的精度分析[J].测绘科学,2021,46(11):12-17. [ Zhao K, Luo L, Yang F Y. Accuracy analysis of precipitable water vapor retrieved by Beidou CORS[J]. Science of Surveying and Mapping, 2021,46(11):12-17. ] DOI:10.16251/j.cnki.1009-2307.2021.11.003
- [11] Xu L, Chen N C, Chen Z Q, et al. Spatiotemporal forecasting in earth system science: Methods, uncertainties, predictability and future directions[J]. Earth-Science Reviews, 2021,222:103828. DOI:10.1016/j.earscirev.2021.103828
- [12] Zhang T, Liu J, Wang J L. Rainstorm prediction via a deep spatio-temporal-attributed affinity network[J]. Geocarto International, 2022,37(26):13079-13097. DOI:10.1080/10106049.2022.2076914
- [13] 程诗奋,彭澎,张恒才,等.异质稀疏分布时空数据插值、重构与预测方法探讨[J].武汉大学学报(信息科学版),2020,45(12):1919-1929. [ Cheng S F, Peng P, Zhang H C, et al. Review of interpolation, reconstruction and prediction methods for heterogeneous and sparsely distributed geospatial data[J]. Geomatics and Information Science of Wuhan University, 2020,45(12):1919-1929. ] DOI:10.13203/j.whugis20200488
- [14] Zheng Z D, Su D C. Short-term traffic volume forecasting: A k-nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm [J]. Transportation Research Part C: Emerging Technologies, 2014,43:143-157. DOI:10.1016/j.trc.2014.02.009
- [15] Zhao L, Song Y J, Zhang C, et al. T-GCN: A temporal graph convolutional network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020,21(9):3848-3858. DOI:10.1109/TITS.2019.2935152
- [16] Chen X Y, Sun L J. Bayesian temporal factorization for multidimensional time series prediction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022,44(9):4659-4673. DOI:10.1109/TPAMI.2021.3066551
- [17] Xu Y, Han L Z, Zhu T Y, et al. Generic Dynamic Graph Convolutional Network for traffic flow forecasting[J]. Information Fusion, 2023,100:101946. DOI:10.1016/j.inffus.2023.101946
- [18] Janowicz K, Gao S, McKenzie G, et al. GeoAI: Spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond[J]. International Journal of Geographical Information Science, 2020,34(4):625-636. DOI:10.1080/13658816.2019.1684500
- [19] 耿佳辰,沈石,程昌秀.“十三五”时期黄河流域PM<sub>2.5</sub>时空分布规律及多尺度社会经济影响机制分析[J].地球信息科学学报,2022,24(6):1163-1175. [ Geng J C, Shen S, Cheng C X. Spatio-temporal evolution and the multi-

- scale socio-economic influencing mechanism of PM<sub>2.5</sub> in the Yellow River Basin during the China's 13<sup>th</sup> five-year plan[J]. Journal of Geo-information Science, 2022,24(6): 1163-1175. ] DOI:10.12082/dqxxkx.2022.210534
- [20] 谢晓苇,李代超,卢嘉奇,等.基于移动监测数据的不同城市场景下PM<sub>2.5</sub>浓度精细模拟与时空特征解析[J].地球信息科学学报,2022,24(8):1459-1474. [ Xie X W, Li D C, Lu J Q, et al. Fine Simulation and Analysis of Temporal and Spatial Characteristics of PM<sub>2.5</sub> Concentration Distribution in Different Urban Scenarios based on Mobile Monitoring Data[J]. Journal of Geo-information Science, 2022,24(8): 1459-1474. ] DOI:10.12082/dqxxkx.2022.210824
- [21] 张永生,张振超,童晓冲,等.地理空间智能研究进展和面临的若干挑战[J].测绘学报,2021,50(9):1137-1146. [ Zhang Y S, Zhang Z C, Tong X C, et al. Progress and challenges of geospatial artificial intelligence[J]. Acta Geodaetica et Cartographica Sinica, 2021,50(9):1137-1146. ] DOI:10.11947/j.AGCS.2021.2020042
- [22] 张彤,刘仁宇,王培晓,等.感知物理先验的机器学习及其在地理空间智能中的研究前景[J].地球信息科学学报,2023,25(7):1297-1311. [ Zhang T, Liu R Y, Wang P X, et al. Physics-informed machine learning and its research prospects in GeoAI[J]. Journal of Geo-information Science, 2023,25(7):1297-1311. ] DOI:10.12082/dqxxkx.2023.2.20795
- [23] Chen R T Q, Rubanova Y, Bettencourt J, et al. Neural ordinary differential equations[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. ACM, 2018:6572-6583. DOI:10.5555/3327757.3327764
- [24] Ji J H, Wang J Y, Jiang Z, et al. STDEN: Towards physics-guided neural networks for traffic flow prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022,36(4):4048- 4056. DOI:10.1609/aaai.v36i4.20322
- [25] Huang C. STR-ODEs: Spatial-Temporal-Ridership graph ODEs for metro ridership prediction[M]. arXiv, 2021. <http://arxiv.org/abs/2107.04980>. DOI:10.48550/arXiv.2107.04980
- [26] Rubanova Y, Chen R T Q, Duvenaud D. Latent ODEs for irregularly-sampled time series[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2019: 5320-5330. DOI:10.48550/arXiv.1907.03907
- [27] Wang P X, Zhang T, Zhang H C, et al. Adding attention to the neural ordinary differential equation for spatio-temporal prediction[J]. International Journal of Geographical Information Science, 2024,38(1):156- 181. DOI:10.1080/13658816.2023.2275160
- [28] Lipton Z C, Berkowitz J, Elkan C. A critical review of recurrent neural networks for sequence learning[M]. arXiv, 2015. <http://arxiv.org/abs/1506.00019>. DOI:10.48550/arXiv.1506.00019
- [29] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016:770-778. DOI:10.1109/CVPR.2016.0
- [30] Zheng Y, Yi X, Li M, et al. Forecasting fine-grained air quality based on big data[C]//Proceedings of the 21th SIGKDD conference on Knowledge Discovery and Data Mining. KDD 2015. <https://www.microsoft.com/en-us/research/publication/forecasting-fine-grained-air-quality-based-on-big-data/>.

■ 本文图文责任编辑:蒋树芳 黄光玉